Journal of Epidemiology and Biostatistics (2001) Vol. 6, No. 6, 417–424

Using splines to analyse latency in the Colorado Plateau uranium miners cohort

M HAUPTMANN[1], K BERHANE[2], B LANGHOLZ[2] and JH LUBIN[1]
[1]*Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda MD, USA*
[2]*Department of Preventive Medicine, University of Southern California, School of Medicine, Los Angeles CA, USA*

Background Different approaches have been proposed to investigate latency in epidemiologic studies where detailed exposure histories are available.

Methods We demonstrate the application of a flexible, yet parsimonious, spline function model to investigate latency patterns for radon progeny exposure and lung cancer in the Colorado Plateau uranium miners cohort. The model extends a previously proposed bilinear model.

Results The excess relative risk (ERR) reached a maximum of 0.6 per 100 working level months, for exposures received 14 years previously. The ERR then declined, and was estimated to approach zero for exposures received 35 years and more in the past. The pointwise 95% confidence intervals supported ERRs > 0 for the period 9–32 years before the event. The estimated latency curve was homogeneous across categories of attained age, duration of exposure, rate of exposure, and smoking.

Conclusions The proposed spline model is a flexible tool for latency analyses, and extends previously used methods.

Keywords lung cancer, radon, nested case-control, latency, spline.

Introduction

Disease latency refers to the interval between an increment of exposure and a subsequent change in an individual's risk. This implies that the risk from a certain exposure history does not depend on cumulative exposure alone, but also on the timing of exposure. It can be expected that risk varies smoothly over time, and this variation can be described by a latency curve.

We applied a spline function model to data from a cohort study of Colorado Plateau uranium miners, to investigate the relation between occupational radon progeny exposure and lung cancer. Splines are piecewise polynomial functions and have been described earlier by de Boor[1]. Theoretically, latency patterns could be described by estimating separate risk parameters for exposures received in each year prior to current age. However, the number of parameters would be large, and all the parameters could not be estimated, due to limited data and high correlations. Spline functions are used to reduce the large number of parameters that would have to be estimated in such a nonparametric approach. Cubic splines induce only mild restrictions, while retaining great flexibility for approximating smooth functions. The model that includes total cumulative exposure, which corresponds to a constant latency curve, is nested within the spline formulation. Other approaches to analyse latency in epidemiologic studies include simple exploratory techniques[2,3] or the bilinear model developed by Langholz *et al.*[4]

We briefly describe the spline function model and apply it to the Colorado Plateau uranium miners cohort. The results are compared with those obtained by Langholz *et al.*[4] using the bilinear model.

Materials and methods

The Colorado Plateau uranium miners cohort study

The Colorado Plateau uranium miners cohort was assembled to study the effeccts of occupational exposure to radioactive radon gas and its progeny, and smoking, on lung cancer mortality. The cohort has been described in detail[4] and consists of male miners recruited between 1950 and 1960. The cohort was traced for vital status through to the end of 1990. We limited our analysis to white males.

We used the nested case-control data described by Langholz *et al.*[4] This data set was drawn from the cohort of 2704 miners, including 263 lung-cancer deaths. For each lung-cancer death, 40 controls (or fewer if necessary) were randomly selected from those who were in the study at the age of death of the case and had attained the age of death of the case during the same 5-year calendar period. Subjects were allowed to serve as controls for more than one case. The analysis data set consisted of 10 322 individuals, including 263 cases.

Correspondence to: M. Hauptmann, National Cancer Institute, 6120 Executive Blvd., Bethesda, MD 20892-7244 USA

Received 9 May 2001
Revised 9 July 2001
Accepted 27 August 2001

Annual exposure to radon and its decay products in working level months (WLM) was estimated based on measurements taken or estimated in the mines. One working level (WL) equals any combination of radon progeny in 1L of air that results in the ultimate emission of 130 000 MeV of energy from α-particles. WLM is a time-integrated exposure measure and is the product of time in units of working months, which is taken to be 170 h, and WL. No measurements were recorded prior to 1950, so the analysis was restricted to those miners who began working in the mines after this date. Details of the exposure reconstruction are available in a technical report[5]. We did not use exposures in the first and second year prior to the death of the case, or the corresponding age for the control, to avoid including exposures that may have been affected by the individual's disease.

Statistical analysis

Linear excess relative risk (ERR) models were fitted, using conditional likelihood regression. Exposure to radon progeny was included as total cumulative exposure in WLM. A 'piecewise constant model' used cumulative exposure during six time intervals as separate continuous covariates (3–5, 6–10, 11–15, 16–20, 21–30, and 31–40 years ago). Wald confidence intervals (CI) were calculated unless otherwise noted. Nested models were compared by likelihood ratio (LR) tests.

The spline model

Individual exposure histories from age at death or current age, to 40 years prior, were used to estimate the ERR of yearly exposure to radon progeny. Let $x(t)$ be the WLM exposure during the year from t-1 to t years prior to the death of a case, or the corresponding age for a control. For example, for an individual with attained age 52.4, $x(1)$ is the WLM exposure from age 51.4 through 52.4, and so forth. Yearly exposures $x(1)$ and $x(2)$ were set to zero in analyses to account for a 2-year lag period.

Thus, $x(3), \ldots, x(40)$ represent the exposure history, and $\Sigma_{t=3}^{40}\ x(t)$ is cumulative WLM. We start with the general model for the relative risk (RR), RR $= 1 + \Sigma_{t=3}^{40}\ \theta_t x(t)$, where $\theta_3, \ldots, \theta_{40}$ are parameters that fully describe the latency curve. In general, data will be insufficient to estimate the full set of parameters, $\theta_3, \ldots, \theta_{40}$. Our approach is then to apply mild constraints to the θ_t's and estimate a functional form that describes their behaviour. Suppose RR $= 1 + \Sigma_{t=3}^{40}\ s(t;\theta)x(t)$, where $s(t;\theta)$ is a function of time t and a parameter vector θ that models the year-specific ERR per WLM, i.e. $s(t;\theta)$ is the ERR per WLM received t years in the past. The weighted sum $\Sigma_{t=3}^{40}\ s(t;\theta)x(t)$ represents the ERR for the exposure profile $x(3), \ldots, x(40)$ compared to a zero profile, i.e., a non-exposed individual.

A cubic B-spline is used to model $s(t;\theta)$.[6] Splines are smooth (i.e. continuously differentiable) piecewise polynomial functions. They are segmented by interior knots. Cubic splines have certain optimal properties for the approximation of curves[1]. The parameters cannot be interpreted directly, but the estimated spline function and corresponding CIs can be plotted.

Spline models with different number and placement of knots are not nested. The number and placement of knots cannot therefore be evaluated by LR tests. To assure a smooth curve and to avoid over-fitting, we considered cubic splines with a small number of interior knots. Two approaches were applied to determine the placement of knots. A profile likelihood search was performed for one interior knot by evaluating the deviance of models for a series of possible knot locations. This approach is computationally cumbersome for multiple knots. Alternatively, knot positions were selected such that the study population accumulated approximately constant proportions of its cumulative exposure between two adjacent knots. For a cubic spline with one interior knot, five spline parameters have to be estimated and the knot position has to be determined. For details see Appendix.

The simple linear ERR model in cumulative exposure is included in the spline model when the function is constant over time, that is $s(t;\theta) = \beta$ for all t. In this case, β is the ERR per WLM. A LR test was performed to test whether the data were consistent, with no variation in the year-specific risk, i.e., cumulative exposure. The spline model was extended by including parameters γ_i in the model RR $= 1 + \gamma_i\ \Sigma_{t=3}^{40}\ s(t;\theta)x(t)$ to evaluate effect modification for attained age, duration of exposure, rate of exposure, and smoking. This was accomplished by categorising each factor by tertiles of the case distribution. LR tests were used to test homogeneity of the multiplicative shift parameters γ_i across categories of these variables. All models for the evaluation of smoking as a potential effect modifier also included a multiplicative adjustment to the baseline δ_i, RR $= \delta_i[1 + \gamma_i\ \Sigma_{t=3}^{40}\ s(t;\theta)x(t)]$, to allow for a main effect of smoking. Effect modification was further evaluated by allowing for different spline functions for categories of potential effect modifiers, RR $= 1 + \gamma_i\ \Sigma_{t=3}^{40}\ s(t;\theta_i)x(t)$, with different sets of spline parameters θ_i and hence latency curves, for each category i of the potential effect modifier.

Results

Descriptive information for radon progeny exposure and smoking is given in Table 1. Cases had higher total exposure than controls. There was substantial variation in timing of exposure; however, this variation was similar in cases and controls. Mean exposure rate among

Table 1 Descriptive statistics of radon progeny and smoking exposure by cases and controls and age, from the Colorado Plateau uranium minders' data

	Age < 60		Age ≥ 60	
	Controls	**Cases**	**Controls**	**Cases**
Number of records[a]	5360	134	4699	129
Total cumulative WLM 3–40 years of latency[b]				
25% percentile	150	478	160	302
50% percentile	372	946	383	679
75% percentile	814	1957	806	1154
Cumulative radon 3–10 years of latency[b]				
No exposure	72%	63%	86%	82%
Median (WLM) among exposed	143	174	89	101
Cumulative radon 11–20 years of latency[b]				
No exposure	41%	33%	53%	42%
Median (WLM) among exposed	168	437	184	286
Cumulative radon 21–40 years of latency[b]				
No exposure	38%	31%	23%	19%
Median (WLM) among exposed	309	691	323	622
Timing of cumulative exposure[c]				
≥ 90% in 3–20 years of latency	41%	41%	25%	22%
≥ 90% in 21–40 years of latency	42%	40%	59%	58%
< 90% in either	17%	19%	16%	20%
Exposure rate during 3–40 years of latency[b] (WL)[d]				
25% percentile	3.0	6.1	3.2	4.5
50% percentile	5.7	8.8	6.0	6.8
75% percentile	9.8	13.6	10.1	11.2
Smoking				
Nonsmokers	22%	7%	23%	11%
Amount smoked among smokers (100s of packs)				
25% percentile	81	83	104	134
50% percentile	116	113	155	165
75% percentile	141	145	190	208

[a] Subjects may be controls in multiple case-control sets.

[b] Years of latency = years before exit from the cohort.

[c] Based on the percentage of total exposure within the latency period.

[d] Computed as the total exposure during 3–40 years of latency divided by the time exposed.

cases and controls by calendar time is shown for categories of attained age in Figure 1. Cases had higher exposure rates than controls. Yearly mean exposure in controls was generally constant over calendar time and across age groups. Yearly mean exposure in cases was higher after 1960 compared with before 1960 for attained ages less than 60 years, and the reverse for attained ages ≥ 60 years.

Table 2 presents deviances for several models. The linear ERR model using total cumulative WLM resulted in an estimated ERR of 0.28 per 100 WLM (95% LR CI: 0.16, 0.51). The piecewise constant model did not fit

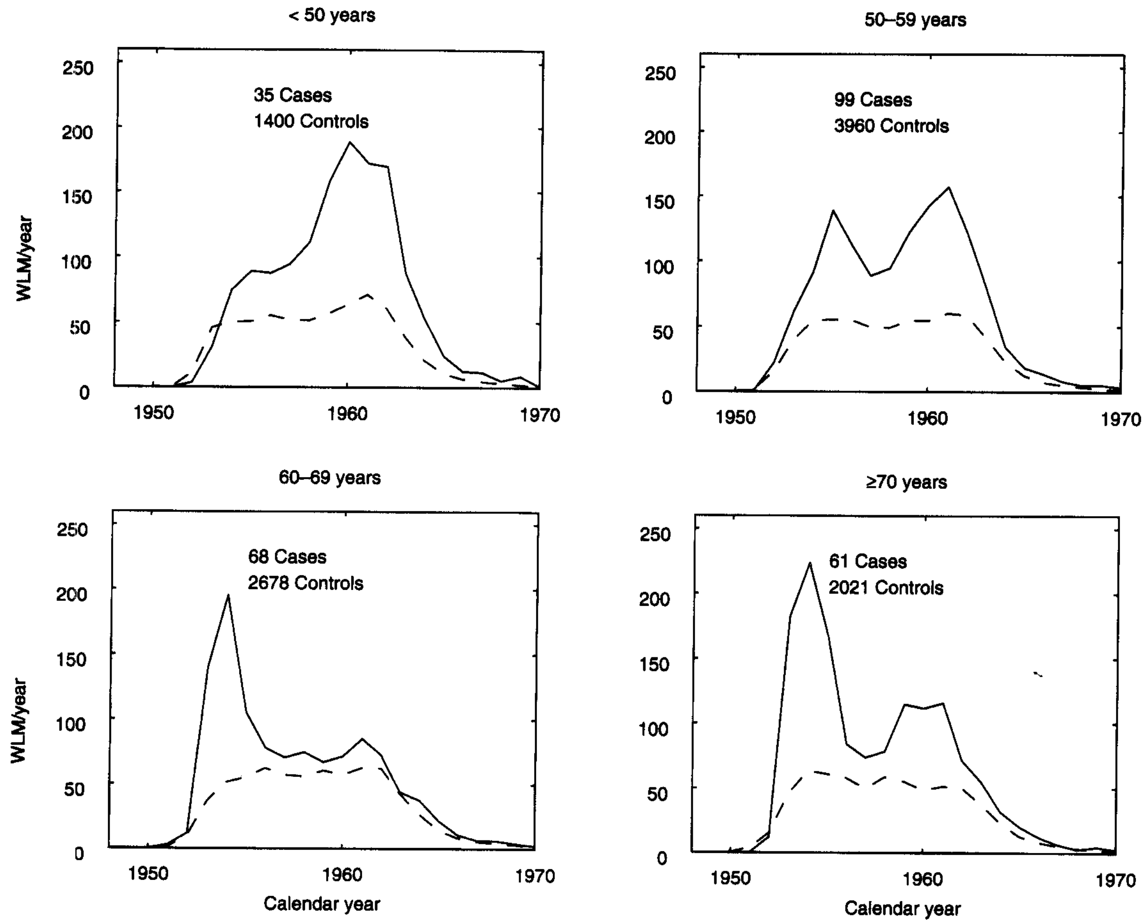

Fig. 1. Yearly mean radon progeny exposure (WLM) among Colorado Plateau uranium miners for cases (solid line) and controls (dashed line) by groups of attained age on calendar-year scale.

significantly better than the total cumulative exposure model (LR test, $p = 0.13$), suggesting no variation in effects with time since exposure. The estimated ERR per 100 WLM were 0.24, 0.44, 0.47, 0.62, 0.30 and 0.08, for exposures received 3–5, 6–10, 11–15, 16–20, 21–30, and 31–40 years ago, respectively. The results are displayed in Figure 2.

For the spline model with one interior knot, the deviance increased with knot location over the entire range from 3 to 40 years prior, with a total change in deviance of 0.5. The maximum likelihood estimate for the interior knot therefore lies on the left boundary of the interval. As a consequence, we fitted a cubic spline function with no interior knot, i.e. a cubic polynomial, to the data. This model fitted the data significantly better than total cumulative exposure ($p = 0.007$, see Table 2).

The spline model with no interior knot is shown in Figure 2. The ERR estimated by the spline function reached a maximum of 0.6 per 100 WLM for exposures

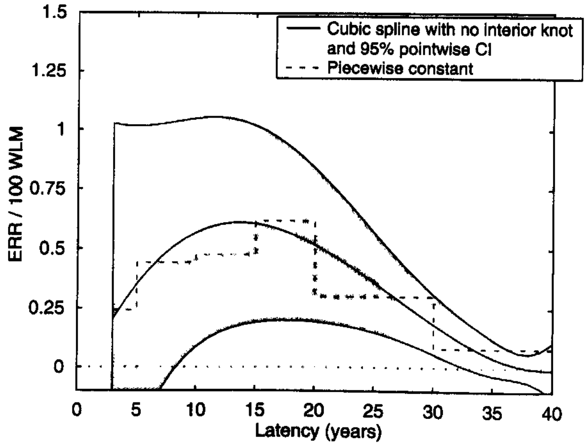

Fig. 2. Excess relative lung cancer risk for the Colorado Plateau uranium miners as a function of latency, based on a piecewise constant model and a cubic spline with no interior knot.

Table 2 Analysis of deviance for comparison of latency models using conditional regression

	Model		Likelihood ratio[a]		
Model	**DF**	**Deviance**	**Statistic**	**DF**	**P**
Cumulative exposure	1	1815.8	–	–	–
Piecewise constant	6	1807.3	8.5	5	0.13
Cubic spline					
No interior knot	4	1803.7	12.1	3	0.007
1 interior knot at 4 years[b]	5	1803.2	12.6	4	0.013
1 interior knot at 20 years[c]	5	1803.7	12.1	4	0.017
2 interior knots at 15, 24 years[d]	6	1803.6	12.2	5	0.032

[a] Compared with the cumulative exposure model.

[b] Location of knot based on profile likelihood search, estimation of knot location not included in degrees of freedom.

[c] Location of knot based on the median of total cumulative cohort exposure over time.

[d] Location of knot based on tertiles of total cumulative cohort exposure over time.

received 14 years prior, then declined and approached zero at ≥ 35 years before the event. The pointwise 95% CI supported ERRs > 0 for the period 9–32 years before the event. As an example, consider two individuals with the same exposure profile except for a difference of 100 WLM at 20 years of latency. The individual with the higher exposure has an ERR of about 0.5 compared with the other individual. If the difference occurs at different years of latency, the ERR changes. The ERR for an exposure profile compared with a zero profile, i.e. a non-exposed individual, can be obtained from the figure by multiplying each yearly WLM with the ERR per WLM for that year of latency and summing up. Note that the spline function agrees closely with the piecewise constant function. The likelihood was flat as a function of the position of the single interior knot, i.e. there was no strong need for a single interior knot. However, we explored spline models with one and more interior knots. The knots were placed according to median, tertiles and quartiles of the total cumulative cohort exposure over time, as explained in the Appendix. Table 2 shows the deviances for those models with one and two interior knots. As expected, the fit did not substantially improve with inclusion of additional knots. Figure 3 displays the estimated spline functions with one and two interior knots with pointwise 95% CI. All curves are very similar to the most parsimonious spline model with no interior knot.

Table 3 shows the effects of including multiplicative parameters for various effect modifiers into the cumulative exposure model and the spline model with no interior knot. In the cumulative exposure model, there was

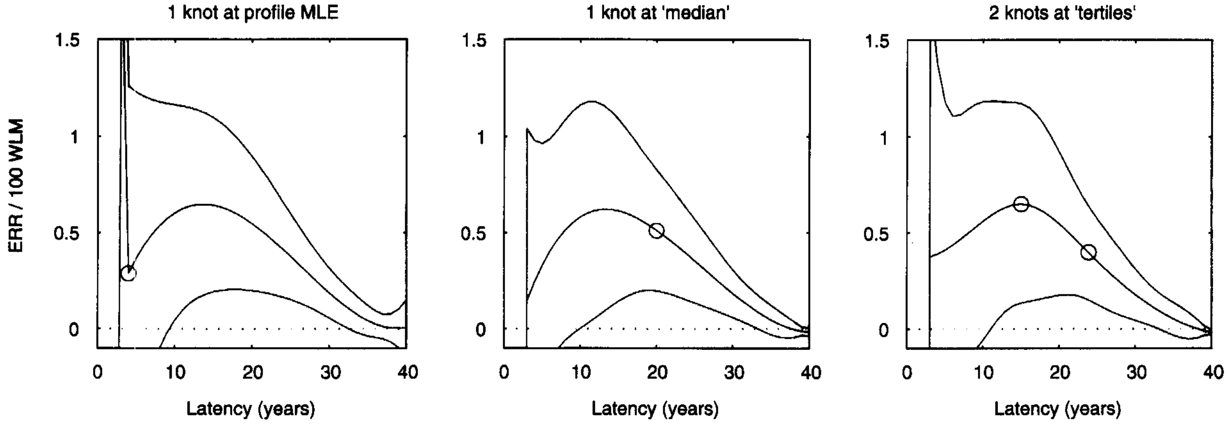

Fig. 3. Excess relative lung cancer risk for the Colorado Plateau uranium miners as a function of latency, based on a cubic spline with different numbers and locations of interior knots (indicated by circles) and pointwise 95% CI (shaded area).

Table 3 Comparison of effect modifying factors for the linear excess relative risk model with cumulative exposure and the cubic spline function with no interior knot

	Cumulative model		Spline model		LR
Modifying variable	**RR**	**Deviance**^{a}	**RR**	**Deviance**^{b}	**P**^{c}
Attained age (years)					
< 55	1.0	1805.6	1.0	1800.8	0.19
55–64	0.7		1.1		
≥ 65	0.1		0.3		
LR P^d	0.006		0.23		
Duration^e (years)					
< 7	1.0	1801.0	1.0	1790.5	0.01
7–10	2.1		2.0		
≥ 11	3.5		3.0		
LR P^d	0.0006		0.001		
Exposure ratee (WL)					
< 6	1.0	1800.1	1.0	1793.2	0.08
6–10	0.9		1.0		
≥ 11	0.4		0.5		
LR P^d	0.0004		0.005		
Smoking (packs)					
< 10 000	1.0	1788.8	1.0	1778.5	0.02
10 000–15 999	0.4		0.6		
≥ 16 000	0.3		0.9		
LR P^d	0.22		0.61		

[a] Deviance of cumulative model without effect modifier is 1815.8 when smoking is not included as a main effect, 1791.8 otherwise

[b] Deviance of cubic spline model without effect modifier is 1803.7, when smoking is not included as a main effect, 1779.5 otherwise

[c] Likelihood ratio test of goodness of fit: cumulative model including effect modifier vs. spline model including effect modifier

[d] Likelihood ratio test of homogeneity

[e] Based on 3 to 40 years prior to the death of the case or the corresponding age for the control.

strong evidence for effect-modification of cumulative radon progeny exposure by categories of attained age, duration and exposure rate ($p \leq 0.006$), but not for smoking ($p = 0.22$). This is consistent with other analyses on the Colorado data[7]. There is no evidence in the spline model for heterogeneity of the multiplicative shift with attained age ($p = 0.23$); for duration and exposure rate homogeneity is still rejected ($p = 0.001$ and $p = 0.005$, respectively). In evaluating the shape of the latency curve, tests of homogeneity of the spline function parameters across categories of attained age, duration, exposure rate, and smoking were not significant ($p > 0.4$, not shown).

Table 3 also shows a comparison of the spline function model with the cumulative model, after allowing for different multiplicative effect modifier parameters. The spline model did not fit significantly better than the cumulative exposure model when a multiplicative shift parameter for attained age ($p = 0.19$) or exposure rate ($p = 0.08$) was included in both models. However, the spline model did fit significantly better than the cumulative exposure model when a multiplicative shift parameter for duration ($p = 0.01$) or smoking ($p = 0.02$) was included.

Discussion

The analysis of latency in epidemiologic studies requires exposure histories that vary over time. The descriptive statistics, as well as the graphical display of yearly mean radon progeny levels, showed substantial variation in timing of exposure in this data set.

The order of the spline and the number and placement of the knots cannot be easily determined 'automatically'. To produce the most parsimonious smooth latency curve, we recommend a cubic spline with a small number of knots to allow adequate flexibility and yet avoid over-fitting. This could be done through an exploratory grid search. For testing, especially for differences in latency shape over levels of an effect modification variable, we suggest minimising the number of knots to the number that well represents the curve. From our experience, there was rarely a need to consider more than three knots.

Fitting cubic spline models with multiple knots did not show a strong decrease in deviance in the Colorado data. We also conducted a sensitivity analysis to examine whether the cubic spline model was sensitive to our choice of maximum latency. Cubic spline weight functions were nearly identical for models with latencies of 30, 35 or 40 years.

It has to be noted that there were some numerical problems during the course of the analysis. It was not always possible to fit models that allow for different sets of spline parameters for categories of effect modifier variables, especially for splines with two or more interior knots. These convergence problems may have been due to the small number of cases compared to the large number of degrees of freedom. The latency curves that we were able to fit for categories of effect modifier variables (with one or no interior knot) tended to exhibit erratic tail behaviours, probably due to scarcity of data at the boundaries. We would like to point out that the boundaries of these latency curves are estimated with less precision (and hence have wider confidence limits). Further research is needed to explore splines that are constrained to be less flexible at the data boundaries.

Langholz *et al.*[4] applied a bilinear model to the same data set. Instead of a cubic spline function they used a more restrictive bilinear function. Their model was characterised by three time-points on the latency scale. There is no effect from exposures received up to time α_0, i.e., ERR $= 0$ for $t \leq \alpha_0$. The effect then increases linearly with time prior to the event, reaching a peak α_1 years in the past, and decreases linearly thereafter, reaching zero at α_2 years in the past, i.e. ERR $= 0$ for $t \geq \alpha_2$. The bilinear model is included in the general class of spline models, namely as a subset of linear splines with two interior knots and with the constraint that the spline function is everywhere non-negative. However, the parametrisation chosen by Langholz *et al.*[4] is different from our framework and requires more restrictive assumptions.

An evaluation of the latency function by categories of several variables showed no evidence for heterogeneity of the latency curve with attained age, duration of exposure, rate of exposure, and smoking. For age, this is contrary to the results obtained by Langholz *et al.*[4], who found different latency patterns for individuals younger than 60 years versus individuals older than 60 years. In their subsequent analysis of four age groups, the peak was estimated to be larger and earlier for younger subjects (ERR 0.74–1.02 per 100 WLM around 10 years in the past) than for older subjects (ERR 0.10–0.14 per 100 WLM around 20 years in the past). The linear slope of the decrease of risk was approximately the same for all age groups (ERR decreased by about 0.02 per year after the peak), so that risk reached the background level earlier for older subjects (around 30 years in the past) compared with younger subjects (around 50 years in the past). However, these estimates were unstable because of the relatively small number of cases in each age group. The fact that there was no evidence for heterogeneity in the spline function may have been due to the greater flexibility of the spline as compared with the bilinear function, to the larger number of degress of freedom in the spline tests, or to the overall better fit for the spline model.

Uncertainties in exposure assessment may become especially apparent when detailed exposure histories are used. It can be expected that the exposure levels reported for the miners are less accurate farther into the past because measurements typically became more frequent and systematic with time. Langholz *et al.*[4] applied the bilinear model to measurement-error-adjusted exposure data generated by Stram *et al.*[8]. The latency curve parameters did not change much compared with the original, unadjusted exposure data. This suggests that, in this cohort, the estimated latency pattern is not affected by measurement error. In summary, the spline function is a flexible yet parsimonious approach for the investigation of latency patterns, as demonstrated with radon progeny exposure and lung cancer in the Colorado Plateau uranium miners cohort.

References

1 de Boor C. *A practical guide to splines*. Number 27 in Applied Mathematical Science. New York: Springer, 1978.

2 Finkelstein MM. Use of time windows to investigate lung cancer latency intervals at an Ontario steel plant. *Am J Indust Med* 1991;19:229–35.

3 Hauptmann M, Lubin JH, Rosenberg PS *et al.* The use of sliding time windows for the exploratory analysis of temporal effects of smoking histories in lung cancer. *Stat Med* 2000; 19:2185–94.

4 Langholz B, Thomas DC, Xiang A, Stram D. Latency analysis in epidemiologic studies of occupational exposures: applications to the Colorado Plateau uranium miners cohort. *Am J Indust Med* 1999;35:246–56.

5 Stram D, Langholz B, Thomas DC. *Measurement error correction of lung cancer risk estimates in the Colorado Plateau*

cohort. I. Dosimetry analysis. Division of Biostatistics Technical Report 126. University of Southern California, School of Medicine, Department of Preventive Medicine, 1998.

6 Hauptmann M, Wellmann J, Lubin JH *et al.* The analysis of exposure-time–response relationships using a spline weight function. *Biometrics* 2000;56:1105–8.

7 Lubin JH, Boice JD, Edling C *et al.* Radon and lung cancer risk: a joint analysis of 11 underground miners studies. NIH publication 94-3644, US National Institutes of Health, 1994.

8 Stram DO, Langholz B, Huberman M, Thomas DC. Correcting for exposure measurement error in a reanalysis of lung cancer mortality for the Colorado Plateau uranium miners cohort. *Health Phys* 1999;77:265–75.

9 Preston DL, Lubin JH, Pierce DA, McConney ME. *Epicure Release 2.0.* Seattle: HiroSoft International Corporation, 1996.

Appendix

Spline function estimation

The function $s(t;\theta)$ is modeled using a B-spline, as described by de Boor[1]. A spline of order k on the interval $[a,b]$ consists of polynomials of order k on the $m + 1$ segments defined by m inner knots $a < t_1 < \ldots < t_m < b$. Adjacent polynomials are smoothly joined, so that the polynomials and their first and second derivatives agree at the knots.

Using a numerically favorable representation of splines, the space of splines can be spanned with $m + k$ basis functions $B_i(t)$, called B-splines. The knot list has to be augmented by $2k$ associated arbitrary 'slack' knots. Without loss of generality, let $t_{-(k-1)} = a - (k - 1)$, $t_{-(k-2)} = a - (k - 2), \ldots, t_0 = a$ and $t_{m+1} = b$, $t_{m+2} = b + 1, \ldots, t_{m+k} = b + k - 1$.

Starting with $B_{i,1}(t) = 1$ if $t_i \leq t < t_{i+1}$ and zero otherwise, the B-spline basis functions are defined by the recurrence relation:

$$B_{i,k}(t) = \frac{t - t_i}{t_{i+k-1} - t_i} B_{i,k-1}(t) + \frac{t_{i+k} - t}{t_{i+k} - t_{i+1}} B_{i+1,k-1}(t) \quad (1)$$

The spline function has the form $s(t;\theta) = \Sigma^m{}_{i=-(k-1)} \theta_i B_{i,k}(t)$. Calculations were performed in EPICURE[9]. The spline parameters were estimated by maximising the likelihood function.

When the best location of a single interior knot was estimated by a profile likelihood search, the maximum likelihood estimator was determined by evaluating the likelihood function for the series $a + 1, a + 2, \ldots, b - 1$ of possible locations of the single interior knot. It has to be noted that the pointwise confidence intervals for the estimated spline function do not include the added variability from estimating the knot position (Figure 3, left panel).

Alternatively, for m inner knots and thus $m + 1$ intervals, knot locations were chosen such that each interval included $1/(m + 1) \times 100\%$ of the total cumulative study population exposure. More precisely: the jth knot t_j was chosen so that $t_j = \max \{t = a, \ldots, b \mid \Sigma_{i=1}^{n} \Sigma^t{}_{l=a} x_i(l) / \Sigma_{i=1}^{n} \Sigma^b{}_{l=a} x_i(l) \leq (j - 1)/(m + 1)\}$, where n is the number of subjects in the study. The computer code for all analyses is available from the first author.

For the cubic spline ($k = 4$) with no interior knot ($m = 0$) between $a = 3$ and $b = 40$, four spline parameters had to be estimated. They were $\hat{\theta} = (0.00116, 0.0127, -0.000757, -0.0000468)$ with standard deviations $\hat{\sigma}(\hat{\theta}) = (0.00471, 0.00578, 0.00273, 0.000811)$.